Replica Symmetric Bound for Restricted Isometry Constant

Ayaka Sakata
The Institute of Statistical Mathematics
Tachikawa, Japan
Theoretical Biology Laboratory
RIKEN, Wako, Japan
Email: ayaka@ism.ac.jp

Yoshiyuki Kabashima
Dep. of Computational Intelligence & Systems Science
Tokyo Institute of Technology
Yokohama, Japan
Email: kaba@dis.titech.ac.jp

***Abstract*—We develop a method for evaluating restricted isometry constants (RICs). This evaluation is reduced to the identification of the zero-points of entropy density which is defined for submatrices that are composed of columns selected from a given measurement matrix. Using the replica method developed in statistical mechanics, we assess RICs for Gaussian random matrices under the replica symmetric (RS) assumption. In order to numerically validate the adequacy of our analysis, we employ the exchange Monte Carlo (EMC) method, which has been empirically demonstrated to achieve much higher numerical accuracy than naive Monte Carlo methods. The EMC method suggests that our theoretical estimation of an RIC corresponds to an upper bound that is tighter than in preceding studies. Physical consideration indicates that our assessment of the RIC could be improved by taking into account the replica symmetry breaking.**

I. Introduction

The signal processing paradigm of compressed sensing (CS) enables a substantially more effective sampling than that required by the conventional sampling theorem [1]. CS is applied to problems in various fields, in which the acquisition of data is quite costly, such as astronomical and medical imaging [2], [3]. The CS performance is mathematically analyzed using the problem settings of a randomized linear observation [4], [5]. Here, $\boldsymbol{A}\in\mathbb{R}^{M\times N}$ is the given observation matrix, and CS endeavors to reconstruct the S-sparse signal $\boldsymbol{x}\in\mathbb{R}^N$ that has $S(<N)$ nonzero components from observation $\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$.

A widely used strategy for the reconstruction of this signal is the ℓ_1 minimization, which corresponds to the relaxed problem of ℓ_0 minimization. A key quantity used to analyze the ℓ_0 and ℓ_1 minimization strategies is the restricted isometry constant (RIC) [6]. Literally evaluating an RIC requires the computation of maximum and minimum eigenvalues of $N!/(S!(N-S)!)$ submatrices that are generated by extracting S-columns from $\boldsymbol{A}$, which is computationally infeasible. In the case of Gaussian random matrices of $\boldsymbol{A}$, the upper bound for the RIC is estimated using the large deviation property without direct computation of the eigenvalues [6], [7], [8].

This paper proposes a theoretical scheme for the direct estimation of the RICs. In order to do this, we evaluate the entropy density of the submatrices that provide a given value of the maximum/minimum eigenvalues. An RIC of matrix $\boldsymbol{A}$ is offered by the condition that the corresponding entropy vanishes. Furthermore, in order to demonstrate our method's utility, we apply our scheme to Gaussian random matrices, using the replica method, and compare the obtained result with that of earlier studies.

Our theoretical evaluation is also numerically assessed using the exchange Monte Carlo (EMC) sampling [9], which is expected to achieve much higher numerical accuracy than those of naive Monte Carlo schemes. The EMC method enables effective sampling, avoiding entrapment at local minima, which limits the effectiveness of naive Monte Carlo sampling to capture the true behavior [10]. Numerical results suggest that our scheme currently provides the tightest RIC upper bound, which could be further tightened by taking into account the replica symmetry breaking (RSB).

II. Restricted isometry constant

In the following, we assume that $\forall\boldsymbol{A}\in\mathbb{R}^{M\times N}$ is normalized so as to (typically) satisfy $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})_{ii}=1$ for all $i\in\{1,\cdots,N\}$.

Definition 1 (Restricted isometry constants)**.** *A matrix $\boldsymbol{A}\in\mathbb{R}^{M\times N}$ satisfies the restricted isometry property (RIP) with RIC $0<\delta_S^{\min}\le\delta_S^{\max}$ if*

$$(1-\delta_S^{\min})||\boldsymbol{x}||_F^2\le||\boldsymbol{A}\boldsymbol{x}||_F^2\le(1+\delta_S^{\max})||\boldsymbol{x}||_F^2 \qquad (1)$$

holds for any S-sparse vector $\boldsymbol{x}\in\mathbb{R}^N$, in which S is the number of non-zero components.

The original work presented by Candès et al. [4] addresses symmetric RIC $\delta_S=\max[\delta_S^{\min},\delta_S^{\max}]$. An RIC indicates how close the space, which is spanned by the S-columns of $\boldsymbol{A}$, is to an orthonormal system. If an RIC is small, the linear transformation performed using $\boldsymbol{A}$ is nearly an orthogonal transformation.

The symmetric RIC provides sufficient conditions for the reconstruction of S-sparse vector $\boldsymbol{x}$ in underdetermined linear system $\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$ using ℓ_0 and ℓ_1 minimization [6].

Theorem 1. *Let $\boldsymbol{A}\in\mathbb{R}^{M\times N}$ and $\boldsymbol{x}\in\mathbb{R}^N$ with $M<N$, and consider the linear equation $\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$. If $\delta_{2S}<1$, a unique S-sparse solution exists and is the sparsest solution to ℓ_0 problem*

$$\min_{\boldsymbol{x}}||\boldsymbol{x}||_0,\ \textit{subject to}\ \boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}. \qquad (2)$$

Also, if $\delta_{2S} < \sqrt{2}-1$, the S-sparse solution to ℓ_1 problem

$$\min_{\boldsymbol{x}} ||\boldsymbol{x}||_1, \text{ subject to } \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{3}$$

is uniquely identified as the sparsest solution and equals the ℓ_0 problem's solution.

It should be noted that $\delta_S^{\min}$ and $\delta_S^{\max}$ do not increase or decrease at the same rate, and asymmetric RICs improve the condition of ℓ_1 reconstruction [11].

Theorem 2. *Consider the same problem settings as in Theorem 1. If $(4\sqrt{2}-3)\delta_{2S}^{\min} + \delta_{2S}^{\max} < 4(\sqrt{2}-1)$, then the unique S-sparse solution is the sparsest solution to the ℓ_1 problem and equals the solution to the ℓ_0 problem [11].*

RIC evaluation is also a fundamental linear algebra problem [7], [8] because RICs clearly relate to the eigenvalues of Gram matrices. Let $T \subseteq V = \{1, \cdots, N\}$, $|T| = S$ be the position of the nonzero elements of S-sparse vector $\boldsymbol{x}$. The product $\boldsymbol{A}\boldsymbol{x}$ equals $\boldsymbol{A}_T\boldsymbol{x}_T$, where $\boldsymbol{A}_T$ is the submatrix that consists of $i \in T$ columns of $\boldsymbol{A}$ and where $\boldsymbol{x}_T = \{x_i | i \in T\}$. For any realization of T, the following holds.

$$\lambda_{\min}(\boldsymbol{A}_T^{\mathrm{T}}\boldsymbol{A}_T)||\boldsymbol{x}_T||_F^2 \le ||\boldsymbol{A}_T\boldsymbol{x}_T||_F^2 \le \lambda_{\max}(\boldsymbol{A}_T^{\mathrm{T}}\boldsymbol{A}_T)||\boldsymbol{x}_T||_F^2$$

Here, $\lambda_{\min}(\boldsymbol{B})$ and $\lambda_{\max}(\boldsymbol{B})$ denote the minimum and maximum eigenvalues of $\boldsymbol{B}$, respectively, and superscript T denotes the matrix transpose. Therefore, the following expression of the RIC is equivalent to eq. (1):

$$\delta_S^{\min} = 1 - \lambda_{\min}^*(\boldsymbol{A};S), \quad \delta_S^{\max} = \lambda_{\max}^*(\boldsymbol{A};S) - 1, \tag{4}$$

in which

$$\lambda_{\min}^*(\boldsymbol{A};S) = \min_{T:T\subseteq V, |T|=S} \lambda_{\min}(\boldsymbol{A}_T^{\mathrm{T}}\boldsymbol{A}_T), \tag{5}$$

$$\lambda_{\max}^*(\boldsymbol{A};S) = \max_{T:T\subseteq V, |T|=S} \lambda_{\max}(\boldsymbol{A}_T^{\mathrm{T}}\boldsymbol{A}_T). \tag{6}$$

Literal evaluation of eq. (4) requires the calculations of the maximum and minimum eigenvalues of the $N!/(S!(N-S)!)$ Gram matrices $\{\boldsymbol{A}_T^{\mathrm{T}}\boldsymbol{A}_T\}$, which is computationally difficult when N and S are large. For typical Gaussian random matrices $\boldsymbol{A}$, the RIC's upper bound is estimated using large deviation properties of the maximum and minimum eigenvalues of the Wishart matrix [6], [7], [8].

III. Problem setup and formalism

We estimate RICs in a different manner, and the following theorem is fundamental to our approach.

Theorem 3. *Let $\boldsymbol{A} \in \mathbb{R}^{M\times N}$. Then the minimum and maximum eigenvalues of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ are given by*

$$\lambda_{\min}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) = -\lim_{\beta\to+\infty} \frac{2}{N\beta}\log Z(\boldsymbol{A};\beta), \tag{7}$$

$$\lambda_{\max}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) = -\lim_{\beta\to-\infty} \frac{2}{N\beta}\log Z(\boldsymbol{A};\beta), \tag{8}$$

respectively, where $Z(\boldsymbol{A};\beta)$ is defined using $\boldsymbol{u} \in \mathbb{R}^N$:

$$Z(\boldsymbol{A};\beta) = \int d\boldsymbol{u} e^{\frac{\beta}{2}||\boldsymbol{A}\boldsymbol{u}||_F^2}\delta(||\boldsymbol{u}||_F^2 - N). \tag{9}$$

Proof: Applying identity $\delta(||\boldsymbol{u}||_F^2 - N) = \beta/(4\pi)\times \int d\eta \exp\left(-\beta\eta/2(||\boldsymbol{u}||_F^2 - N)\right)$ gives us

$$Z(\boldsymbol{A};\beta) = \frac{(2\pi)^{\frac{N}{2}-1}}{2\beta^{\frac{N}{2}-1}}\int d\eta \exp\Big[\beta\Big\{\frac{N\eta}{2} - \frac{1}{2\beta}\sum_i \ln(\eta+\lambda_i)\Big\}\Big],$$

in which $\{\lambda_i\}$ is the ith eigenvector of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$. As $\beta \to +\infty$, the integral can be evaluated using the saddle point method, which is dominated by $\eta = -\lambda_{\min}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) + (N\beta)^{-1} + o(\beta^{-1})$, where $o(\beta^{-1})$ represents the contribution from negligible terms compared with β^{-1}. This yields eq. (7), and eq. (8) is similarly obtained by applying the saddle point method for $\beta \to -\infty$. $\square$

Theorem 3 holds for all submatrices $\boldsymbol{A}_T$. For mathematical convenience, we introduce variables $\boldsymbol{c} \in \{0,1\}^N$ and define

$$Z_c(\boldsymbol{c},\boldsymbol{A};\beta) = \int d\boldsymbol{u} P(\boldsymbol{u}|\boldsymbol{c})\exp\Big\{-\frac{\beta}{2}||\boldsymbol{A}(\boldsymbol{c}\circ\boldsymbol{u})||_F^2\Big\} \times \delta(||\boldsymbol{c}\circ\boldsymbol{u}||_F^2 - N), \tag{10}$$

where $\circ$ denotes the component-wise product, and $P(\boldsymbol{u}|\boldsymbol{c}) \propto \exp\Big(-\sum_{i=1}^N (1-c_i)u_i^2/2\Big)$ is introduced in order to avoid the divergence caused by integrating u_i when $c_i = 0$. Let us define $\boldsymbol{c}(T) \in \{0,1\}^N$ to be $(\boldsymbol{c}(T))_i = 1$ for $i \in T$ and to be $(\boldsymbol{c}(T))_i = 0$ otherwise. The two functions $Z(\boldsymbol{A}_T;\beta)$ and $Z_c(\boldsymbol{c}(T),\boldsymbol{A};\beta)$ have a one-to-one correspondence: $Z(\boldsymbol{A}_T;\beta) = Z_c(\boldsymbol{c}(T),\boldsymbol{A};\beta)$. We write $\lambda_{\max}(\boldsymbol{c},\boldsymbol{A})$ and $\lambda_{\min}(\boldsymbol{c},\boldsymbol{A})$, which are obtained by substituting $Z_c(\boldsymbol{c},\boldsymbol{A};\beta)$ into eq. (7) and eq. (8), respectively. Because $\lambda_{\max}(\boldsymbol{A}_T^{\mathrm{T}}\boldsymbol{A}_T) = \lambda_{\max}(\boldsymbol{c}(T),\boldsymbol{A})$ and $\lambda_{\min}(\boldsymbol{A}_T^{\mathrm{T}}\boldsymbol{A}_T) = \lambda_{\min}(\boldsymbol{c}(T),\boldsymbol{A})$ naturally hold, eqs. (5-6) can be respectively rewritten as

$$\lambda_{\min}^*(\boldsymbol{A};S) = \min_{\boldsymbol{c}\in\boldsymbol{c}_S}\lambda_{\min}(\boldsymbol{c},\boldsymbol{A}), \tag{11}$$

$$\lambda_{\max}^*(\boldsymbol{A};S) = \max_{\boldsymbol{c}\in\boldsymbol{c}_S}\lambda_{\max}(\boldsymbol{c},\boldsymbol{A}), \tag{12}$$

where $\boldsymbol{c}_S$ denotes the set of configurations of $\boldsymbol{c}$ that satisfy $\sum_i c_i = S$.

We hereafter focus on the situation in which both M and S are proportional to N as $M = N\alpha$ and $S = N\rho$, respectively, where $\alpha, \rho \sim O(1)$. Let us define the *energy* densities of $\boldsymbol{c}$ to be $\Lambda_+(\boldsymbol{c}|\boldsymbol{A}) = \lambda_{\min}(\boldsymbol{c},\boldsymbol{A})/2$ and $\Lambda_-(\boldsymbol{c}|\boldsymbol{A}) = \lambda_{\max}(\boldsymbol{c},\boldsymbol{A})/2$. Based on this, we introduce a *free entropy* density as $\phi(\mu|\boldsymbol{A};\rho) = N^{-1}\log\Big[\sum_{\boldsymbol{c}} e^{-N\mu\Lambda_{\mathrm{sgn}(\mu)}(\boldsymbol{c}|\boldsymbol{A})}\delta\big(\sum_{i=1}^N c_i - N\rho\big)\Big]$, where $\mathrm{sgn}(\mu)$ denotes the sign of μ. Eqs. (7-8) offer its alternative expression

$$\phi(\mu|\boldsymbol{A};\rho) = \lim_{\frac{\beta}{\mu}\to+\infty} N^{-1}\log\Big[\sum_{\boldsymbol{c}} Z_c^{\frac{\mu}{\beta}}(\boldsymbol{c},\boldsymbol{A};\beta)\delta\Big(\sum_{i=1}^N c_i - N\rho\Big)\Big]. \tag{13}$$

In addition, we represent the number of $\boldsymbol{c}$ that correspond to $\Lambda_\pm(\boldsymbol{c}|\boldsymbol{A}) = \lambda/2$ and satisfy $\sum_i c_i = N\rho$ as $\exp(N\omega_\pm(\lambda|\boldsymbol{A};\rho))$ using *entropy* densities $\omega_\pm(\lambda|\boldsymbol{A};\rho)$, which are naturally assumed to be convex functions of λ. Summation over the microscopic states of $\boldsymbol{c}$ is replaced with the integral of λ over

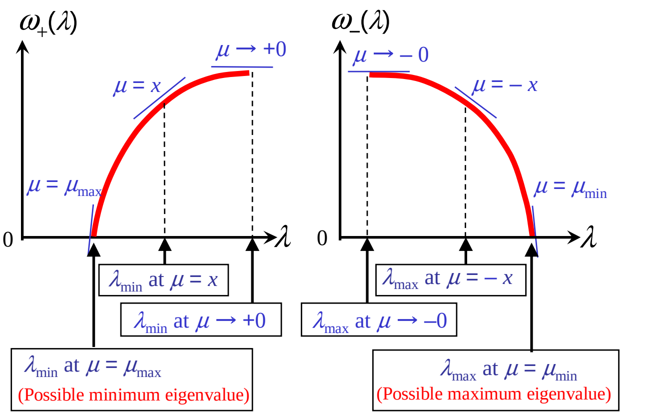

Fig. 1. Schematic picture of entropy curve and relationship to parameter μ.

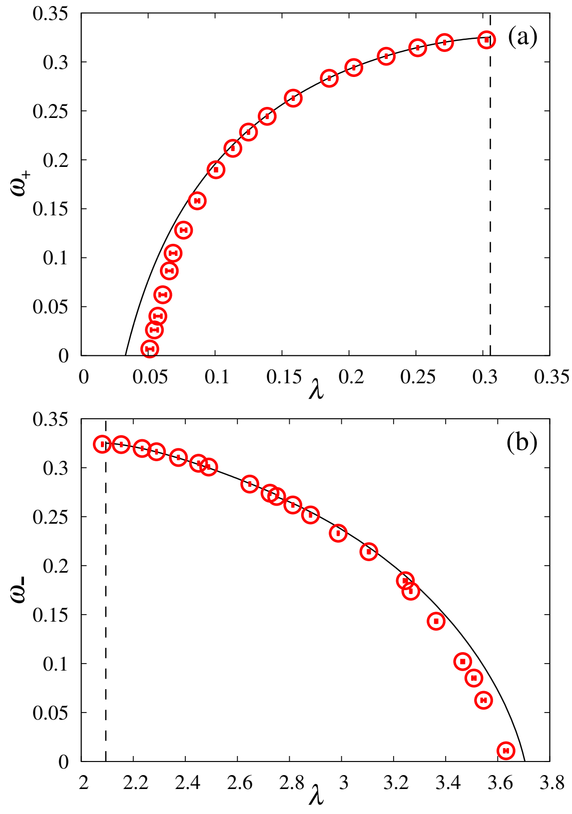

Fig. 2. Entropy curve for $\alpha = 0.5$ and $\rho = 0.1$ with (a) $\mu > 0$ and (b) $\mu < 0$. Circles denote EMC method results. Vertical lines represent (a) minimum and (b) maximum eigenvalues of MP distribution.

the possible value of $\Lambda_{\pm}(\boldsymbol{c}|\boldsymbol{A})$:

$$\phi(\mu|\boldsymbol{A};\rho) = \frac{1}{N}\log\left[\int d\lambda \exp\{-N\mu\lambda + N\omega_{\mathrm{sgn}(\mu)}(\lambda|\boldsymbol{A};\rho)\}\right] \rightarrow \max_{\lambda}\{-\mu\lambda + \omega_{\mathrm{sgn}(\mu)}(\lambda|\boldsymbol{A};\rho)\}, \tag{14}$$

in which the saddle point method is employed. The maximizer of λ, which corresponds to the typical energy value of $\boldsymbol{c}$ that is sampled following the weight $e^{-N\mu\Lambda_{\mathrm{sgn}(\mu)}(\boldsymbol{c}|\boldsymbol{A})}\delta(\sum_{i=1}^{N} c_i - N\rho)$, must satisfy

$$-\mu + \frac{\partial \omega_{\mathrm{sgn}(\mu)}(\lambda|\boldsymbol{A};\rho)}{\partial\lambda} = 0. \tag{15}$$

Eq. (14) implies that $\phi(\mu|\boldsymbol{A};\rho)$ is obtained using the Legendre transformation of $\omega_{\pm}(\lambda|\boldsymbol{A};\rho)$, and the inverse Legendre transformation converts $\phi(\mu|\boldsymbol{A};S)$ to $\omega_{\pm}(\lambda|\boldsymbol{A};\rho)$ as

$$\omega_{\mathrm{sgn}(\mu)}(\lambda|\boldsymbol{A};\rho) = \phi(\mu|\boldsymbol{A};\rho) - \mu\frac{\partial\phi(\mu|\boldsymbol{A};\rho)}{\partial\mu}, \tag{16}$$

from the convexity assumption of $\omega_{\pm}(\lambda|\boldsymbol{A};\rho)$. A similar formalism has been introduced for investigating the geometrical structure of weight space in learning of multilayer neural networks [12].

The relationships among μ, λ, and $\omega_{\pm}$ are illustrated in Fig. 1. Entropy densities ω_+ and ω_- are convex increasing and decreasing functions of λ, respectively. According to eq. (15), the value of λ at μ represents the point where the gradient of $\omega_{\pm}$ equals μ. By definition, negative entropy values are not allowed, and $\omega_{\pm}(\lambda|\boldsymbol{A};\rho) < 0$ implies that no $\boldsymbol{c}$ simultaneously satisfies both $\Lambda_{\pm}(\boldsymbol{c}|\boldsymbol{A}) = \lambda/2$ and $\sum_i c_i = N\rho$. Therefore, the $\lambda_{\pm}^*$ that produces $\omega_{\pm}(\lambda_{\pm}^*|\boldsymbol{A};\rho)$=0 is the possible minimum or maximum eigenvalue. Hence, eqs. (11-12) give us

$$\lambda_{\min}^*(\boldsymbol{A};\rho) = \lambda_+^*, \quad \lambda_{\max}^*(\boldsymbol{A};\rho) = \lambda_-^*, \tag{17}$$

which are the typical values for $\mu = \mu_{\max}$ and $\mu = \mu_{\min}$, respectively (Fig. 1).

IV. RS ANALYSIS FOR GAUSSIAN RANDOM MATRIX

This section applies the methodology introduced in the previous section to the case in which components of $\boldsymbol{A}$ are independently generated using a Gaussian distribution with mean 0 and variance $(N\alpha)^{-1}$. In this case, $\phi(\mu|\boldsymbol{A};\rho)$ and $\omega_{\pm}(\lambda|\boldsymbol{A};\rho)$ randomly fluctuate depending on $\boldsymbol{A}$. However, for all $\varepsilon > 0$, the probability that deviation from the typical values, $\phi(\mu;\rho) \equiv [\phi(\mu|\boldsymbol{A};\rho)]_A$ and $\omega_{\pm}(\lambda;\rho) \equiv [\omega_{\pm}(\lambda|\boldsymbol{A};\rho)]_A$, is larger than ε tends to vanish as $N \to \infty$. Here, $[\cdot]_A$ denotes the average of $\boldsymbol{A}$. Therefore, typical properties can be characterized by evaluating the typical values, $\phi(\mu;\rho)$ and $\omega_{\pm}(\lambda)$, using the replica method with the identity [13], [14]:

$$[\log f(\boldsymbol{A})]_A = \lim_{n\to 0}\frac{\partial}{\partial n}\log[f^n(\boldsymbol{A})]_A \tag{18}$$

where $f(\boldsymbol{A})$ is an arbitrary function. When both n and $m = \mu/\beta$ are positive integers, regarding $\sum_{\boldsymbol{c}} Z_c^m(\boldsymbol{c},\boldsymbol{A};\beta)\delta(\sum_{i=1}^N c_i - N\rho)$ in eq. (13) as $f(\boldsymbol{A})$ leads us to express $[f^n(\boldsymbol{A})]_A$ as a summation/integration with respect to n and nm replica variables $\{\boldsymbol{c}^a\}$ and $\{\boldsymbol{c}^a \circ \boldsymbol{u}^{a\sigma}\}$ $(a \in \{1,2,\ldots,n\}, \sigma \in \{1,2,\ldots,m\})$, which can be evaluated by the saddle point method for $N \to \infty$.

Under the replica symmetric (RS) assumption, in which the dominant saddle point is assumed to be invariant against any permutation of the replica indices a and σ within each of their sets $\{1,2,\ldots,n\}$ and $\{1,2,\ldots,m\}$, respectively, the resulting

functional form of $N^{-1}\log[f^n(\boldsymbol{A})]_A$ becomes extendable for non-integer n and m. Therefore, we insert the expression into eq. (13) employing the formula of eq. (18), which finally yields

$$\phi(\mu;\rho) = -\frac{\alpha}{2}\log\{\alpha+\chi+\mu(1-q)\} + \frac{\alpha}{2}\log(\alpha+\chi)$$
$$-\frac{\alpha\mu q}{2\{\alpha+\chi+\mu(1-q)\}} + \frac{\hat{Q}}{2} - \frac{\hat{q}_1}{2}\left(1+\frac{\chi}{\mu}\right) + \frac{\hat{q}_0 q}{2} + K\rho$$
$$+\int Dz\log\left\{1+e^{-K}\int Dy\exp\left(\frac{(\sqrt{\hat{q}_1-\hat{q}_0}y+\sqrt{\hat{q}_0}z)^2}{2\hat{Q}}\right)\right\}, \quad (19)$$

where $\{q,\chi,\hat{Q},\hat{q}_0,\hat{q}_1,K\}$ are determined to extremize the right hand side, and $\int Dz = \int_{-\infty}^{+\infty}\frac{dz}{\sqrt{2\pi}}\exp(-z^2/2)$. The derivation of eq. (19) is shown in Appendix A. Entropy densities $\omega_\pm(\lambda;\rho)$ are derived by applying the inverse Legendre transformation to $\phi(\mu;\rho)$.

V. Results

In Fig. 2, entropy densities $\omega_\pm$ with $\alpha=0.5$ and $\rho=0.1$ are shown for (a) $\mu<0$ and (b) $\mu>0$. Results of the exchange Monte Carlo (EMC) sampling are represented by circles, and the EMC procedure is summarized in Appendix B.

The values of λ when $\mu\to+0$ and $\mu\to-0$, which are denoted using dashed lines, coincide with the respective minimum and maximum of the Marchenko-Pastur (MP) distribution's support for the $M\times S$ Gaussian random matrix [15]. As the limit of $|\mu|\to 0$ corresponds to unbiased generation of $M\times S$ Gaussian random matrices, the coincidence theoretically supports the adequacy of our analysis. The slight discrepancy between the theoretical and EMC results in the entropy's tails could be due to the insufficiency of the RS assumption. The convexity of our entropy suggests that the RS assumption exactly creates the entropy curve or extends it outward [16]. This is consistent with the EMC method's result, which indicates that the exact entropy curve is inward when compared to that produced by the RS assumption. Therefore, the estimated zero-points, $\lambda^*_{\max}$ and $\lambda^*_{\min}$, that are provided using the RS assumption, are meaningful upper and lower bounds, respectively, of the true values. We call them RS bounds.

Fig. 3 compares our RS upper bound, Bah and Tanner's upper bound [6], and the RIC numerically obtained lower bound [17]. In this example, the symmetric RIC is $\delta_S^{\max}$. Our analysis lowers the upper bound of the RIC, especially for a large ρ/α region. Over the entire parameter region, our estimates are consistent with the numerically obtained lower bound.

Fig. 4 shows the parameter region that mathematically supports ℓ_0 and ℓ_1 reconstruction according to Theorems 1 and 2. The region determined by the Bah and Tanner RIC is indicated using black lines. The RS bound of the RIC extends the region in which correct reconstruction is guaranteed, and further extension may be provided by taking the RSB into account.

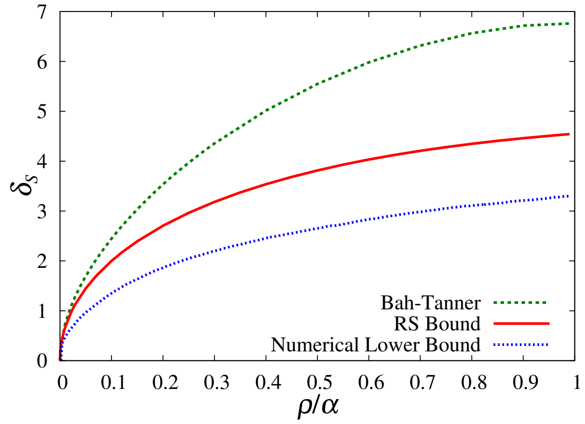

Fig. 3. Comparison of symmetric RIC for $\alpha=0.5$. Numerical lower bound is estimated for $N=1000$ and $M=500$.

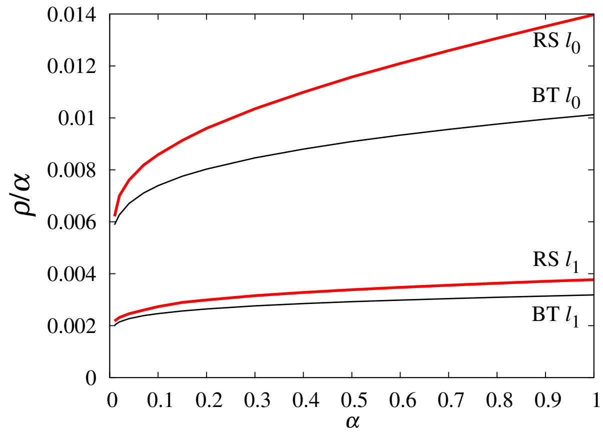

Fig. 4. ℓ_0 and ℓ_1 limits given by RS RIC. Black lines represent Bah and Tanner's results, denoted by BT.

VI. Summary and Conclusion

We proposed a theoretical scheme for the evaluation of restricted isometry constants. The problem was converted to the assessment of entropy density, and the possible maximum and minimum eigenvalues, which produce the RIC, are the entropy's zero-points. Given a Gaussian random matrix, we computed the entropy density using the replica method under the replica symmetric ansatz and estimated the value of the RIC. Physically, it has meaning as a bound and is tighter than existing bounds. Numerical experiments using the EMC sampling support our analysis.

A more accurate evaluation of the RIC is possible if the RSB is taken into account. Our scheme is applicable to more general matrices than Gaussian random matrices as well.

Appendix A
RS calculation of free entropy density

Identities

$$1 = \int dq^{(a,\sigma)(b,\tau)}\delta\left(q^{(a,\sigma)(b,\tau)} - \frac{1}{N}\sum_{i=1}^{N} c_i^a c_i^b u_i^{a\sigma} u_i^{b\tau}\right), \quad (20)$$

for all combinations of replica indices (a,σ) and (b,τ) $(a,b=1,2,\dots,n;\sigma,\tau=1,2,\dots,m)$, are employed in the saddle point assessment of $\phi_\beta(n,m;\rho)\equiv N^{-1}\log[(\sum_{\boldsymbol{c}} Z^m(\boldsymbol{c},\boldsymbol{A};\beta)\delta(\sum_{i=1}^N c_i-N\rho))^n]_A$. We assume that the dominant saddle point is of the replica symmetric form as

$$q^{(a,\sigma)(b,\tau)}=\begin{cases}1 & \text{for } a=b,\ \sigma=\tau\\ q_1 & \text{for } a=b,\ \sigma\neq\tau\\ q_0 & \text{for } a\neq b.\end{cases} \tag{21}$$

This means that when $\boldsymbol{A}$ is a Gaussian random matrix of mean 0 and variance $(N\alpha)^{-1}$,

$$[s_\mu^{a\sigma}s_\nu^{b\tau}]_A=\alpha^{-1}\delta_{\mu\nu}(\delta_{ab}\delta_{\sigma\tau}+q_1\delta_{ab}(1-\delta_{\sigma\tau})+q_0(1-\delta_{ab}))$$

holds, where $s_\mu^{a\sigma}\equiv\sum_i A_{\mu i}c_i^a u_i^{a\sigma}$. Higher order correlations are negligible due to the central limit theorem, which indicates that $s_\mu^{a\sigma}$ can be expressed as $s_\mu^{a\sigma}=\alpha^{-1/2}(\sqrt{1-q_1}w_\mu^{a\sigma}+\sqrt{q_1-q_0}v_\mu^a+\sqrt{q_0}z_\mu)$, where $w_\mu^{a\sigma}$, v_μ^a, and z_μ are i.i.d. Gaussian random variables of zero mean and unit variance. Replacing $[\cdot]_A$ with average with respect to these Gaussain variables, the saddle point evaluation offers an expression of $\phi_\beta(m,\rho)\equiv\lim_{n\to0}\frac{\partial}{\partial n}\phi_\beta(n,m;\rho)$, as

$$\begin{aligned}\phi_\beta(m;\rho)=&\frac{m(\tilde{Q}+\tilde{q}_1q_1)-m^2(\tilde{q}_1q_1-\tilde{q}_0q_0)}{2}+\int Dz\log\Xi_\beta(z)\\&+\alpha\Big[-\frac{m}{2}\log\Big(1+\frac{\beta(1-q_1)}{\alpha}\Big)-\frac{1}{2}\log\Big(1+\frac{\mu(q_1-q_0)}{\alpha+\beta(1-q_1)}\Big)\\&-\frac{\mu q_0}{2\{\alpha+\beta(1-q_1)+\mu(q_1-q_0)\}}\Big]+K\rho,\end{aligned} \tag{22}$$

where

$$\Xi_\beta(z)=1+\frac{e^{-K}}{(\tilde{Q}+\tilde{q}_1)^{m/2}}\int Dy\exp\Big(\frac{m(\sqrt{\tilde{q}_1-\tilde{q}_0}y+\sqrt{\tilde{q}_0}z)^2}{2(\tilde{Q}+\tilde{q}_1)}\Big)$$

and $\tilde{Q}$, K, $\tilde{q}_1$ and $\tilde{q}_0$ are conjugate variables for the integral representations of delta functions in eq. (10), eq. (13) and eq. (20), respectively. Eq. (22) yields the free entropy density as $\phi(\mu;\rho)=\lim_{\beta\to\infty}\phi_\beta(\mu/\beta,\rho)$, in which the variables scale so that $\hat{Q}\equiv m(\tilde{Q}+\tilde{q}_1)$, $\hat{q}_1\equiv m^2\tilde{q}_1$, $\hat{q}_0\equiv m^2\tilde{q}_0$, and $\chi\equiv\beta(1-q_1)$ become $O(1)$. This gives the expression of eq. (19).

The variables $\{\chi,q_0,\hat{Q},\hat{q}_1,\hat{q}_0,K\}$ are determined by extremization conditions of the free entropy density eq. (19),

$$\chi=\frac{\mu\rho}{\hat{Q}} \tag{23}$$

$$q=\int Dz\Big\{\frac{\Xi(z)-1}{\Xi(z)}\frac{\sqrt{\hat{q}_0}z}{\hat{Q}-\hat{\Delta}}\Big\}^2 \tag{24}$$

$$1=\int Dz\frac{\Xi(z)-1}{\Xi(z)}\Big(\frac{\hat{\Delta}}{\hat{Q}(\hat{Q}-\hat{\Delta})}+\frac{\hat{q}_0z^2}{(\hat{Q}-\hat{\Delta})^2}\Big) \tag{25}$$

$$\rho=\int Dz\frac{\Xi(z)-1}{\Xi(z)} \tag{26}$$

$$\hat{\Delta}=\frac{\alpha\mu^2(1-q)}{(\alpha+\chi)\{\alpha+\chi+\mu(1-q)\}} \tag{27}$$

$$\hat{q}_0=\frac{\alpha\mu^2q}{\{\alpha+\chi+\mu(1-q)\}^2} \tag{28}$$

where $q=q_0$, $\hat{\Delta}=\hat{q}_1-\hat{q}_0$, and $\Xi(z)=\lim_{\beta\to\infty}\Xi_\beta(z)$.

Appendix B
Monte Carlo sampling for RIC estimation

We employ the exchange Monte Carlo (EMC) sampling [9] in order to numerically compute the free entropy density $\phi(\mu|\boldsymbol{A};\rho)$ and obtain the entropy density avoiding the trap of metastable states. In the EMC approach, we prepare k systems, which have the same configuration of $\boldsymbol{A}$, and assign configuration $\boldsymbol{c}_i\in\boldsymbol{c}_S$ and parameter μ_i to each system $i=1,\cdots,k$. The signs of $\{\mu_i\}$ are set to be the same. Each step of the EMC process updates $\boldsymbol{c}_i$ within each system, and attempts exchanges between configurations $\boldsymbol{c}_i$ and $\boldsymbol{c}_{i+1}$. The probability of transition from $\boldsymbol{c}_i$ to $\boldsymbol{c}_i'$ is given by $w(\boldsymbol{c}_i,\boldsymbol{c}_i')=\min\{\exp(\mu_iN\Delta_i),1\}$, where $\Delta_i=\Lambda_{\mathrm{sgn}(\mu_i)}(\boldsymbol{c}_i|\boldsymbol{A})-\Lambda_{\mathrm{sgn}(\mu_i)}(\boldsymbol{c}_i'|\boldsymbol{A})$. The probability of an exchange between systems $\boldsymbol{c}_i$ and $\boldsymbol{c}_{i+1}$ is given by $w_{\mathrm{exc}}(\boldsymbol{c}_i,\boldsymbol{c}_{i+1})=\min\{\exp(N(\mu_i-\mu_{i+1})\Delta_{i,i+1}),1\}$, in which $\Delta_{i,i+1}=\Lambda_{\mathrm{sgn}(\mu_i)}(\boldsymbol{c}_i|\boldsymbol{A})-\Lambda_{\mathrm{sgn}(\mu_{i+1})}(\boldsymbol{c}_{i+1}|\boldsymbol{A})$. After sufficient updates, the entire k-system is expected to converge to equilibrium distribution $P_{\mathrm{tot}}(\{\boldsymbol{c}_i\}|\boldsymbol{A})\propto\prod_{i=1}^k\exp\{-N\mu_i\Lambda_{\mathrm{sgn}(\mu_i)}(\boldsymbol{c}_i|\boldsymbol{A})\}$.

The density of states $W_\pm(\lambda|\boldsymbol{A};\rho)=\sum_{\boldsymbol{c}\in\boldsymbol{c}_S}\delta(\lambda/2-\Lambda_\pm(\boldsymbol{c}|\boldsymbol{A}))$ is obtained by applying the multihistogram method [18] using histgrams of $\Lambda_\pm$ obtained by EMC sampling. Finally, the free entropy density is calculated:

$$\phi(\mu|\boldsymbol{A};\rho)=\int d\lambda W_{\mathrm{sgn}(\mu)}(\lambda|\boldsymbol{A};\rho)\exp(-N\mu\lambda/2), \tag{29}$$

and the entropy density is derived by applying the inverse Legendre transformation to eq. (29).

Acknowledgment

The authors would like to thank Tomoyuki Obuchi for his helpful comments and discussions. This work was partially supported by the RIKEN SPDR fellowship and by KAKENHI No. 26880028 (AS), and KAKENHI No. 25120013 (YK).

References

[1] H. Nyquist, Trans. AIEE **47**, 617 (1928).
[2] D. D. Lustig and J. Pauly, Magnet. Reson. Med. **58** (6), 1182–1195 (2007).
[3] J. Bobin et al., IEEE J. Sel. Top. Signal Process. **2** (5), 718–726 (2008).
[4] E. Candès et al., IEEE Trans. Inform. Theory **52** (2), 489–509 (2006).
[5] D. Donoho, IEEE Trans. Inform. Theory **52** (4), 1289–1306 (2006).
[6] E. Candès and T. Tao, IEEE Trans. Inform. Theory **51** (12), 4203–4215 (2005).
[7] J. D. Blanchard et al., SIAM Rev. **53** (1), 105–125 (2011).
[8] B. Bah, and J. Tanner, SIAM J. Matrix Anal. & Appl. **31** (5), 2882–2898 (2010).
[9] K. Hukushima, and K. Nemoto, J. Phys. Soc. Jpn, **65**, 1604–1608 (1996).
[10] D. Donoho and Y. Tsaig, Signal Process. **86** (3), 549–581 (2006).
[11] S. Foucart and M.-J. Lai, Appl. Comput. Harmon. Anal. **26** (3), 395–407 (2009).
[12] R. Monasson and D. O'Kane, Europhys. Lett. **27** (2), 85–90 (1994).
[13] M. Mézard et al., *Spin Glass Theory and Beyond*, (World Sci. Pub., 1987).
[14] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction*, (Oxford Univ. Pr., 2001).
[15] V. A. Marchenko and L. A. Pastur, Mat. Sb. (N. S.) **72** (114), 507–536 (1967).
[16] T. Obuchi et al., J. Phys. A: Math. Theor. **43**, 485004 (28pp) (2010).
[17] C. Dossal et al., Linear Algebra Appl. **432**, 1663–1679 (2010).
[18] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **64**, 1195–1198 (1989).